Linear transvection groups

Hans Cuypers       Anja Steinbach

1 Introduction

Most classical groups arising from (anti-) hermitian forms or (pseudo-) quadratic forms contain so-called isotropic transvections. Indeed, suppose, for example, that $V$ is a vector space over some skew field $K$ endowed with a $(\sigma,-1)$-hermitian form and let $w$ be an isotropic vector of $V$, i.e., $f(w,w) = 0$, with $w \notin \text{Rad}(V,f)$. Then, for each $a \in K^*$ with $a^\sigma = a$, the map $v \mapsto v + f(v,w)aw$ for $v \in V$ is a transvection fixing the form $f$.

The isotropic transvection subgroups of these classical groups, i.e., the subgroups generated by all isotropic transvections with a fixed axis, form a class $\Sigma$ of abelian subgroups which is a class of abstract transvection groups in the sense of Timmesfeld [25]. This means that for all $A,B \in \Sigma$ we have that $[A,B] = 1$ or $\langle A,B \rangle$ is a rank 1 group (i.e., $A \neq B$, and for each $a \in A^\#$, there exists some $b \in B^\#$ with $A^b = B^a$).

Here we describe a common characterization of all these classical groups with isotropic transvections as linear groups generated by a class $\Sigma$ of abstract transvection subgroups such that the elements of $A \in \Sigma$ act as transvections.

Details and proofs of the results mentioned in this paper can be found in [8] and will appear elsewhere.

2 Transvections

2.1 Notation. Suppose $V$ is a left vector space of arbitrary dimension defined over some skew field $K$. For any linear map $t : V \rightarrow V$ (acting from the right) and vector $v \in V$, we define the commutator $[v,t]$ to equal $vt - v$. Here $vt$ is the image of $v$ under the linear map $t$. If $W$ is a subspace of $V$ and $S$ a set of linear maps of $V$, then $[W,S]$ is the subspace of $V$ spanned by $\{[v,s] \mid w \in W, s \in S\}$.

An invertible linear map $t : V \rightarrow V$ is called a transvection if

(a) $[V,t]$ is 1-dimensional and
(b) \([V, t] \subseteq C_Y(t) = \{v \in V \mid vt = v\}\).

Suppose \(t : V \rightarrow V\) is a transvection. From the definition it is clear that \(C_Y(t)\) is a hyperplane of \(V\), it is called the axis of \(t\). The 1-dimensional subspace \([V, t]\) is called the center of \(t\).

Let \(v_t\) be a vector spanning the center \([V, t]\) of \(t\). Then there is an element \(\varphi \in V^*\), the dual of \(V\), with kernel \(C_Y(t)\) such that the action of \(t\) on \(V\) can be described as follows:
\[
t : v \mapsto v + (v \varphi)v_t \text{ for } v \in V.
\]

### 2.2 Transvections in classical groups.

In this paper we consider subgroups of the general linear group on \(V\) which are generated by transvections. For finite-dimensional \(V\), it is well known that the special linear group \(\text{SL}(V)\) is generated by its transvections. For infinite-dimensional \(V\), the subgroup of \(\text{GL}(V)\) generated by the transvections is finitary, i.e., for each element \(g\) of this subgroup the commutator \([V, g]\) is finite-dimensional. In fact the transvections generate the full finitary special linear group \(\text{FSL}(V)\).

Also the classical subgroups of \((\text{F})\text{SL}(V)\) arising from (anti-) hermitian forms contain transvections. Indeed, suppose \(V\) is endowed with a \((\sigma, -1)\)-hermitian form \(f\) and let \(w\) be an isotropic vector of \(V\), i.e., \(f(w, w) = 0\), with \(w \not\in \text{Rad}(V, f)\). Then, for each \(a \in K^*\) with \(a^\sigma = a\), the map
\[
t : v \mapsto v + f(v, w)aw \text{ for } v \in V
\]
is a transvection fixing the form \(f\). Such a transvection will be called isotropic with respect to \(f\), cf. Hahn and O'Meara [10, p. 213]. If the dimension of \(V\) is finite, then the subgroup of \(\text{SL}(V)\) of isometries of \(f\) is generated by its isotropic transvections. Similarly, for infinite-dimensional \(V\), the isotropic transvections leaving \(f\) invariant generate the finitary subgroup of the corresponding classical group.

Maybe less well-known is the following class of transvections which we find in orthogonal groups. Suppose \(\sigma\) is an involutory anti-automorphism of \(K\) and for \(\epsilon \in \{-1, 1\}\), set \(\Lambda := \\{c - cc^\sigma \mid c \in K\}\). Now consider a non-degenerate pseudo-quadratic form \(q : V \rightarrow K/\Lambda\) with associated trace-valued \((\sigma, \epsilon)\)-hermitian form \(f : V \times V \rightarrow K\), see Tits [26, (8.2.1)] (a radical of \(f\) is allowed). Let \(w\) be an isotropic vector of \(V\), i.e., \(q(w) = 0 + \Lambda\). If there exist \(a \in K^*\) and \(r_a \in \text{Rad}(V, f)\) (possibly 0) with \(q(r_a) = a + \Lambda\), then the map
\[
t : v \mapsto v + f(v, w)(aw + r_a) \text{ for } v \in V
\]
is a transvection in the isometry group of \( q \), which we also call an \textit{isotropic transvection}. The axis of \( t \) is the space \( w^\perp = \{ v \in V \mid f(v, w) = 0 \} \), its center is \( \langle aw + r_a \rangle \). Such transvections exist provided that \( q \) is not an ordinary quadratic form with trivial radical \( \text{Rad}(V, f) \).

We notice that these isotropic transvections act trivially on \( \text{Rad}(V, f) \) and therefore also induce transvections on the space \( V/\text{Rad}(V, f) \).

2.3 \textbf{Transvection subgroups}. Let \( t_1 \) and \( t_2 \) be two transvections on \( V \). Up to symmetry we only have the following three possibilities for the centers and axes of \( t_1 \) and \( t_2 \):

1. \( [V, t_1] \subseteq C_V(t_2) \) and \( [V, t_2] \subseteq C_V(t_1) \), then \( [t_1, t_2] = 1 \),

2. \( [V, t_1] \not\subseteq C_V(t_2) \) and \( [V, t_2] \not\subseteq C_V(t_1) \), then \( \langle t_1, t_2 \rangle \) is contained in the group \( \text{SL}([V, t_1] \oplus [V, t_2]) \simeq \text{SL}_2(K) \),

3. \( [V, t_1] \subseteq C_V(t_2) \) and \( [V, t_2] \not\subseteq C_V(t_1) \), then \( [t_1, t_2] \) is also a transvection on \( V \) with center \( [V, t_1] \) and axis \( C_V(t_2) \).

If \( t_1 \) and \( t_2 \) are isotropic transvections with respect to some anti-hermitian form \( f \), or some pseudo-quadratic form \( q \) with associated \( (\sigma, \epsilon) \)-hermitian form \( f \), then case (3) does not occur. Indeed, for all \( v, w \in V \) we have that \( f(v, w) = 0 \) if and only if \( f(w, v) = 0 \).

Now suppose \( t_1 \) and \( t_2 \) are two isotropic transvections with respect to some anti-hermitian form \( f \) or pseudo-quadratic form \( q \). Denote by \( T_1 \) and \( T_2 \), respectively, the subgroup of \( \text{GL}(V) \) generated by all isotropic transvections with the same axis as \( t_1 \) or \( t_2 \), respectively. These subgroups are called \textit{isotropic transvection subgroups} and are isomorphic to \( (K^\sigma, +) \), if we consider the isotropic transvections with respect to the anti-hermitian form, and to \( (\Delta, +) \) in case they leave the pseudo-quadratic form \( q \) invariant. Here \( K^\sigma = \{ a \in K \mid a^\sigma = a \} \) and \( \Delta = \{ a \in L \mid \text{there exists } r_a \in \text{Rad}(V, f) \text{ with } q(r_a) = a + \Lambda \} \).

It is straightforward to check that for \( T_1 \) and \( T_2 \) we have one of the following two possibilities:

1. \( [V, T_1] \subseteq C_V(T_2) \) and \( [V, T_2] \subseteq C_V(T_1) \), then \( [T_1, T_2] = 1 \).

2. \( [V, T_1] \not\subseteq C_V(T_2) \) and \( [V, T_2] \not\subseteq C_V(T_1) \), then \( \langle T_1, T_2 \rangle \) is isomorphic to the subgroup

\[
\langle \left( \begin{array}{cc} 1 & 0 \\ \lambda & 1 \end{array} \right), \left( \begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array} \right) \mid \lambda, \mu \in K^\sigma \text{ resp. } \Delta \rangle
\]

of \( \text{SL}_2(K) \), denoted by \( \text{SL}_2(K^\sigma) \) or \( \text{SL}_2(\Delta) \).
The subgroups $\mathrm{SL}_2(K^\sigma)$ and $\mathrm{SL}_2(\Delta)$ of $\mathrm{SL}_2(K)$ are rank 1 groups in the following sense, see Timmesfeld [25]: for each $x_1 \in T_1^\#$ there exists an $x_2 \in T_2^\#$ with $T_2^{x_1} = T_1^{x_2}$. Indeed, if $1 \neq x_1 = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}$, then for $x_2$ we may take $\begin{pmatrix} 1 & \lambda^{-1} \\ 0 & 1 \end{pmatrix}$. Note that if $\lambda \neq 0$ is in $K^\sigma$ or $\Delta$, respectively, then so is $\lambda^{-1}$.

3 The main results

The main goal is to give a common characterization of the various classical groups (different from the special linear group) as groups generated by their isotropic transvection subgroups. We now describe the exact setting we will work in:

3.1 Setting. Let $K$ be a skew field and $V$ a vector space over $K$. Assume that $G \leq \mathrm{GL}(V)$ such that:

1. $G$ is generated by a conjugacy class $\Sigma$ of abelian subgroups of $G$.

2. For $A, B \in \Sigma$, either $[A, B] = 1$ or $\langle A, B \rangle$ is a rank 1 group (i.e., $A \neq B$, and for each $a \in A^\#$, there exists some $b \in B^\#$ with $A^b = B^a$).

3. For $A \in \Sigma$, every $a \in A^\#$ is a transvection on $V$.

4. Each $A \in \Sigma$ contains at least 3 elements.

5. There are $A, B \in \Sigma$ with $[A, B] = 1$ and $C_\Sigma(A) \neq C_\Sigma(B)$.


The conditions (1) and (2) on $\Sigma$ are the defining conditions of a class of abstract transvection groups in a group $G$ in the sense of Timmesfeld [25]. In (5), we use the definition $C_\Sigma(A) = \{ T \in \Sigma \mid [A, T] = 1 \}$, for $A \in \Sigma$. By $P(V)$ we denote the projective space corresponding to $V$.

Notice that we do not assume that for each $A \in \Sigma$ the commutator space $[V, A]$ is 1-dimensional nor that $C_V(A)$ is a hyperplane in $V$.

We are now able to state our first result:

3.2 Theorem. Assume $G$ is a subgroup of $\mathrm{GL}(V)$ generated by a class $\Sigma$ of abstract transvection groups as in the setting above.

If $C_V(G) = 0$ (e.g., $V$ is irreducible), then $G$ is quasi-simple and we are in one of the Cases (a) to (c):
(a) There exist a skew field $L$ with involutory anti-automorphism $\sigma$, some $\epsilon \in \{1,-1\}$ and a vector space $W$ over $L$ endowed with one of the following forms (recall that $\Lambda := \{c - \epsilon c^\sigma \mid c \in L\}$):

1. a non-degenerate pseudo-quadratic form $q : W \to L/\Lambda$ with associated trace-valued $(\sigma, \epsilon)$-hermitian form $f : W \times W \to L$ or

2. a non-degenerate $(\sigma, \epsilon)$-hermitian form $f$, where $\Lambda$ coincides with $
\{c \in L \mid \epsilon cc^\sigma = -c\}$.

The group $G$ is isomorphic to the classical normal subgroup of the isometry group of $q$ or $f$ generated by the isotropic transvection subgroups $T_p$, where $p$ runs over the points of the classical polar space $Q$ arising from $W$ and $q$ or $f$ (i.e., $p$ runs over the isotropic points of $P(W)$).

(b) There exists a quaternion skew field $L$ (with standard (anti-)involution $\sigma$ and center $Z(L)$). Denote by $W$ the vector space $L^4$ over $L$ endowed with the pseudo-quadratic form $q : W \to L/Z(L)$ defined by $q(x_1, x_2, x_3, x_4) := x_1x_3^\sigma + x_2x_4^\sigma + Z(L)$ for $x_1, x_2, x_3, x_4 \in L$ (with associated $(\sigma, -1)$ hermitian form $f$).

The group $G$ is isomorphic to the subgroup of the isometry group of $q$ generated by those isotropic transvection subgroups $T_p$, where $p$ runs over the points of a so-called special subquadrangle $Q$ of the classical generalized quadrangle arising from $W$ and $q$.

(c) There exists a non-perfect commutative field $L$ of characteristic 2 with

$$L^2 \subseteq \Theta' \subseteq L' \subseteq \Theta \subseteq L,$$

where $L'$ is a subfield of $L$, $\Theta$ is an $L'$-subspace of $L$ which generates $L$ as a ring and $\Theta'$ is an $L^2$-subspace of $L'$ which generates $L'$ as a ring. Denote by $W$ the vector space $\Theta \times (L')^4$ over $L'$ endowed with the quadratic form $q : W \to L'$ defined by $q(x_0; (x_1, x_2, x_3, x_4)) := x_0^2 + x_1x_2 + x_3x_4$ for $x_0 \in \Theta$ and $x_1, x_2, x_3, x_4 \in L'$ (with associated symmetric bilinear form $f$).

The group $G$ is isomorphic to the subgroup of the isometry group of $q$ generated by those isotropic transvection subgroups $T_p$, where $p$ runs over the points of a so-called mixed subquadrangle $Q$ of the classical generalized quadrangle arising from $W$ and $q$.

In Cases (a) to (c), the elements of $\Sigma$ are contained in the isotropic transvection subgroups and in 1-1-correspondence with the points of $Q$.

The action of $G$ on $V$ is induced by a semi-linear mapping $\varphi : W \to V$ (with respect to an embedding $\alpha : L \to K$ resp. $\alpha : L' \to K$ in Case (c)) with kernel Rad$(W, f)$. 
A detailed description of the generalized quadrangles in Cases (b) and (c) can be found in Steinbach and Van Maldeghem [21].

In Theorem 3.2, we may conclude that any two transvections \( t_1 \) and \( t_2 \) in some element \( A \in \Sigma \) have the same center and the same axis. However, in the examples that we have encountered in 2.2 we have seen that \([V, A] \) need not always be a point of \( P(V) \). However, the condition \( C_V(G) = 0 \) forces \([V, A] \) to be 1-dimensional for all \( A \in \Sigma \).

If we relax on the condition \( C_V(G) = 0 \) but still insist on \([V, A] \) being 1-dimensional, we find that most of the groups \( G \) are classical groups, however, now defined by a possibly degenerate form:

\textbf{3.3 Theorem.} Assume \( G \) is a subgroup of \( \mathrm{GL}(V) \) generated by a class \( \Sigma \) of abstract transvection groups as in the setting 3.1 above.

Assume in addition that \([V, A] \) is 1-dimensional for all \( A \in \Sigma \). Then there exists a subspace \( U \) of \( V \) and a subgroup \( G_0 = \langle \Sigma_0 \rangle \), \( \Sigma_0 = \{ A \in \Sigma \mid [V, A] \subseteq U \} \) of \( G \) such that:

\begin{enumerate}
\item[(1)] \( V = U + C_V(G) \) (not necessarily a direct sum);
\item[(2)] \( U \) and \( G_0 \) are as \( V \) and \( G \) in the conclusion of Theorem 3.2 with the possible exception that the kernel of the semi-linear mapping \( \varphi \) is only contained in \( \text{Rad}(W, f) \).
\end{enumerate}

Moreover, if \( U \) and \( G_0 \) are as in Case (a) of 3.2, then \( G \) is isomorphic to the classical group generated by the isotropic transvection subgroups on a vector space \( \overline{W} = W \oplus \tilde{R} \) over \( L \), with extended form \( q \) or \( f \) such that \( \tilde{R} \) is isotropic and contained in \( \text{Rad}(\overline{W}, f) \).

The action of \( G \) on \( V \) is induced by a semi-linear mapping \( \overline{\varphi} : \overline{W} \rightarrow V \) extending \( \varphi \).

In the situation of Theorem 3.3, the classical group generated by the isotropic transvection subgroups on the vector space \( \overline{W} = W \oplus \tilde{R} \) is isomorphic to \( (\oplus_{i \in I} W_i) : G_0 \), where each \( W_i \) is a copy of the natural module \( W \) for \( G_0 \).

We say two elements \( A, B \in \Sigma \) are equivalent if \( C_\Sigma(A) = C_\Sigma(B) \). For each \( A \in \Sigma \), the subspace \( U \) of \( V \) occurring in Theorem 3.3 contains exactly one of the \([V, T] \) where \( T \) runs through the equivalence class of \( A \). In the case where all \([V, T] \) are 1-dimensional and the equality \( C_\Sigma(A) = C_\Sigma(B) \) implies that \( A = B \), we see that \( U = V \) and \( G = G_0 \). Hence this assumption may replace the one that \( C_V(G) = 0 \) in Theorem 3.2 (except for the statement on the kernel of \( \varphi \)).

We may overcome the assumption in Theorem 3.3 that \([V, A] \) is 1-dimensional as follows:
3.4 Proposition. In the setting of 3.1 the subspace $R = \bigcap_{A \in \Sigma} [V, A]$ is contained in $C_V(G)$. Moreover, the codimension of $R$ in $[V, A]$, $A \in \Sigma$, is one, so that we may apply Theorem 3.3 to $G/N \leq \text{GL}(V/R)$, where $N$ is the kernel of the action of $G$ on $V/R$.

The proofs of these results are mainly geometric. They can be found in [8]. As may be clear from the examples appearing in the conclusion of the theorems, all groups act as automorphism groups on a polar space. The main idea in the proof of Theorem 3.2 is to construct this polar space together with an embedding into the projective space $P(V)$ of $V$. Once this is done, we can apply the full strength of the theory of polar spaces and its embeddings. In particular, we use the classification of non-degenerate Moufang polar spaces due to Tits and the classification of their weak embeddings (of degree $> 2$) by Steinbach and Van Maldeghem [18], [21]. We find the isomorphism type of the group $G$ in Theorem 3.2 by identifying $G$ as a group of automorphisms of a polar space. The action of $G$ on the space $V$ in 3.2 is determined by the embedding of this polar space in $P(V)$.

We note that rather than proving that the polar space constructed in Timmesfeld [25] is weakly embedded in $P(V)$ (which is not obvious), we preferred to construct a polar space which is automatically weakly embedded. (Hypothesis (II) of [25] is satisfied in our setting only as long as elements of $\Sigma$ contain at least 4 elements.) The construction does not rely on finite dimensions, commutative fields or perfect fields in characteristic 2. It is a uniform approach resulting in all different types of classical groups.

Theorem 3.2 is an intermediate result in the proof of 3.3. To prove 3.3 we show that there is a subspace $U$ of $V$ such that $V = U + C_V(G)$ and that the centers of abstract transvection groups in $\Sigma$ which are contained in $U$ form a non-degenerate polar space weakly embedded in $P(U)$. We are then able to identify the subgroup $G_0$ of $G$ generated by those elements of $\Sigma$ that have their center in $U$ as a (quasi-simple) classical group generated by the isotropic transvection subgroups. The group $G$ is finally identified as a split extension of $G_0$ by a normal subgroup which is a direct sum of natural modules for $G_0$ (or equivalently, as a classical group arising from a degenerate form).

4 3-Transpositions

If $\Sigma$ is a set of abstract transvection groups of $G$ with $|A| = 2$, for $A \in \Sigma$, then the set of all non-trivial elements in the members of $\Sigma$ is a set of 3-transpositions of $G$. 
Finite groups generated by 3-transpositions have been studied by Fischer, who classified the finite almost simple groups generated by 3-transpositions, see [9]. Recently, Cuypers and Hall [6] gave a complete classification of the centerfree 3-transposition groups containing at least 2 commuting 3-transpositions.

Among the 3-transposition groups we find various examples in which the 3-transpositions are in fact transvections on some natural module. Indeed, the finitary symplectic groups \( \text{FSp}(V, f) \), where \((V, f)\) is a symplectic \( \text{GF}(2) \)-space, and finitary unitary groups \( \text{FSU}(W, h) \), with \((W, h)\) a hermitian space over \( \text{GF}(4) \), are generated by their isotropic transvections which form a class of 3-transpositions.

There are more 3-transposition groups where 3-transpositions are transvections. The group \( \text{Sp}_{2n}(2) \) contains three classes of irreducible subgroups generated by transvections: the symmetric group \( S_{2n+2} \) and the two different types of orthogonal groups, \( O^+_{2n}(2) \) and \( O^-_{2n}(2) \).

The symplectic and unitary groups fit perfectly in the scheme of this paper. With some extra effort we could have extended our methods to include these groups in Theorem 3.2. Indeed, the set of centers of the isotropic transvections in these groups carries the structure of a polar space, see Cuypers [7]. Reconstruction of this polar space, cp. Cuypers [7, Section 3] and Hall [11, Section 2], and its embedding would yield a result similar to Theorem 3.2. The symmetric and orthogonal groups, however, do not fit into our scheme.

5 Some historical remarks.

The roots of the present paper can be found in the work of McLaughlin on linear groups generated by transvections [15] and the work of Fischer on 3-transposition groups, see [9]. Indeed, the study of linear groups generated by transvections has been initiated by the work of McLaughlin who classified all irreducible subgroups of \( \text{GL}(V) \), \( V \) finite-dimensional over a field \( k \), generated by full linear transvection subgroups. (A full linear transvection subgroup of \( \text{GL}(V) \) consists of all transvections to a fixed center and axis in \( \text{GL}(V) \).) Soon after McLaughlin’s work Piper [16], [17], Wagner [28] and Kantor [12] considered subgroups of finite linear groups generated by transvections. They obtained results similar to Theorem 3.2. Where Piper’s and Wagner’s approach was very geometric, Kantor’s work had a more group theoretic flavor. He used the work of Fischer and generalizations thereof by Aschbacher [1], Aschbacher and Hall [2], and Timmesfeld [22].

In more recent years McLaughlin’s work has been extended to greater classes of groups. Li [13] and Vavilov [27] have generalized McLaughlin’s results to finite-
dimensional vector spaces defined over arbitrary skew fields and Cameron and Hall [3] also considered linear transvection groups acting (possibly reducibly) on a module $V$ of arbitrary dimension. Related results can also be found in Cuypers [4] and Timmesfeld [23].

Timmesfeld generalized the concept of 3-transpositions to that of abstract transvection groups. In [24], [25], he obtained a classification of the quasi-simple groups generated by a class of abstract transvection groups, under some additional richness assumption (which implies that the abstract transvection groups contain at least 4 elements). The cases where the abstract transvection groups contain less than 4 elements have been dealt with by Cuypers [5] (3 elements) and by Cuypers and Hall [6] (3-transpositions case). Timmesfeld’s results formed the basis for the work of the second author on subgroups of classical groups generated by long root elements, see Steinbach [19], [20]. With regards to linear groups generated by transvections, she determines the modules $V$ (finite-dimensional over an arbitrary commutative field) for the various groups of Timmesfeld’s classification such that the abstract transvection groups are parts of the linear transvection subgroups of $\text{GL}(V)$.

Liebeck and Seitz [14] classified the closed subgroups of groups of Lie type over algebraically closed fields which are generated by root elements.

The present approach (see [8]) combines both the work on linear groups generated by transvections as begun by McLaughlin and the geometric study of abstract groups initiated by the work of Fischer on 3-transpositions to obtain a common characterization of all the classical groups generated by isotropic transvection subgroups. Although we do not make use of the various results quoted above, several ideas and methods used in this paper come from this work. In particular, the proofs of our main results are obtained by combining the group theoretic methods of Timmesfeld [25], the more geometric approach of Cuypers [5] and the concept of weak embeddings as in Steinbach [19].

References


Hans Cuypers
Department of Mathematics
Eindhoven University of Technology
P.O. BOX 513
5600 MB Eindhoven
The Netherlands
e-mail: hansc@win.tue.nl

Anja Steinbach
Mathematisches Institut
Justus-Liebig-Universität Gießen
Arndtstraße 2
D 35392 Gießen
Germany
e-mail: Anja.Steinbach@math.uni-giessen.de